

Searching by Document Number

\*\* Result [Patent] \*\* Format(P801) 21.Feb.2002 1/ 1  
Application no/date: 1983- 25079[1983/02/16]  
Date of request, for examination: [1989/11/06]  
Public disclosure no/date: 1984-149398[1984/08/27]  
Examined publication no/date (old law): 1991- 58520[1991/09/05]  
Registration no/date: 1694217[1992/09/17]  
Examined publication date (present law): [ ]  
PCT application no: [ ]  
PCT publication no/date: [ ]  
Applicant: SANYO ELECTRIC CO LTD  
Inventor: SHIMIZU MASAHIKA  
IPC: G10L 9/14  
Expanded classification: 425  
Fixed keyword: R108  
Title of invention: VOICE SYNTHESIZATION SYSTEM  
Abstract:

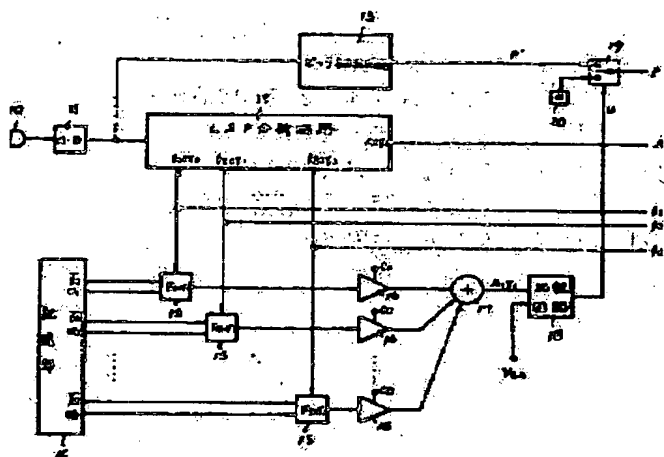
PURPOSE: A sharp reduction of a calculation quantity for judgement processing is planned by judging \*\* silence audio by using the value that consists of the LSP frequency parameters that are drawn by means of executing a specific mathematical-statistical technique in the LSP speech synthesis method.

CONSTITUTION: By memory part 14, normalized circuit 15, multiplier 16 and adder 17//Each mean value  $f_1^* - f_2^*$  of the LSP parameters  $f_1$ ,  $f_2$  and  $f_n^*$ ,  $f_n^*$  and its standard deviations  $\sigma_1$ ,  $\sigma_2$  and  $\sigma_n$  are determined (\*\*\*) . After normalizing by translating  $f_1^*$  and  $f_n^*$  into expression \*1 respectively  $f_1 T$  or  $f_2 T$  of a LSP parameter of which time is  $T$ //In addition, weighting coefficients  $C_1$  and  $C_2$ ,  $\sigma_1$  and  $C_n$  are set, and expression \*2 is calculated. It is output from the LSP analysis circuit 12. Comparison circuit 18//AT is compared with specific threshold value  $V_{th}$ . When it is a silent audio at the time of  $AT > V_{th}$ , it is judged. A silent audio detection signal is output. When it is voiced sound voice at the time of  $AT < V_{th}$ , it is judged to the contrary, and a voiced sound voice detection signal is output.

( Automatic Translation )

$$F_1(z) = \frac{F_1(z) - F_1}{0.1}$$

$$A(z) = \sum_{i=0}^{\infty} 0.1 F_1(z)$$



Registration number(1694217) has already removed to closed files.

Other Translation

⑬ 日本国特許庁 (JP)

⑪ 特許出願公開

⑫ 公開特許公報 (A)

昭59-149398

⑭ Int. Cl.<sup>3</sup>

識別記号

庁内整理番号

⑬ 公開 昭和59年(1984) 8月27日

G 10 L 1/00

7350-5D

発明の数 1  
審査請求 未請求

(全 5 頁)

⑭ 音声合成方式

⑯ 特 願 昭58-25079

⑰ 出 願 昭58(1983) 2月16日

⑱ 発 明 者 清水雅久

守口市京阪本通2丁目18番地三  
洋電機株式会社内

⑲ 出 願 人 三洋電機株式会社

守口市京阪本通2丁目18番地

⑳ 代 理 人 弁理士 佐野静夫

明 細 書

1. 発明の名称

音 声 合 成 方 式

2. 特許請求の範囲

有声音声に対して周期パルス音源信号を発生する有声音源発生回路と、無声音声に対して雑音音源信号を発生する無声音源発生回路と、これ等両音源発生回路の内の一方の回路から得られる音源信号を濾波するデジタルフィルタとを備え、該デジタルフィルタの濾波特性をN個の線スペクトル対パラメータ $f_1, f_2, \dots, f_N$ にて制御する事に依つて、音声波形を合成する音声合成方式に於いて、線スペクトル対パラメータ $f_1, f_2, \dots, f_N$ の各平均値 $\bar{f}_1, \bar{f}_2, \dots, \bar{f}_N$ 及び標準偏差 $\sigma_1, \sigma_2, \dots, \sigma_N$ を予め求めておき、時刻Tの線スペクトル対パラメータ $f_1(T), f_2(T), \dots, f_N(T)$ を夫々

$$F_1(T) = \frac{f_1(T) - \bar{f}_1}{\sigma_1}$$

( $1=1, 2, \dots, N$ )

に変換する事に依つて正規化した後、さらに累み

付け係数 $\alpha_1, \alpha_2, \dots, \alpha_N$ を設定して

$$A(T) = \frac{1}{N} \sum_{i=1}^N \alpha_i F_i(T) \quad (1=1, 2, \dots, N)$$

を算出し、

この値A(T)が特定のしきい値 $V$ より大なる時には、上記無声音源発生回路からの雑音音源信号を上記デジタルフィルタに供給し、逆にこの値A(T)が特定のしきい値 $V$ より小なる時には上記有声音源発生回路からの周期パルス音源信号を上記デジタルフィルタに供給する事を特徴とした音声合成方式。

8. 発明の詳細な説明

H1 産業上の利用分野

本発明は音声の特徴パラメータに基づいて音声を合成する音声合成方式に関する。

H2 従来技術

この種音声合成方式は人間の発声器官の機能を電氣的に模製したものであり、従来のボコーダ技術に始まり、L P O (線形予測符号化) 方式、P A R O O R (部分自己相関) 方式、L B P (線ス

ベクトル対)方式が次々と開発されてきた。中でも最近の開発に依るLBP方式は音声の特徴パラメータとして用いるLBP周波数パラメータの情報圧縮率が極めて高く、音質の高い合成音声を得られる利点を備えているので、現在のPAROOR方式に代つて、音声合成方式に代つて、音声合成方式の主流になりつつある。

第1図にLBP方式を採用した音声合成装置の基本構成を示す。同図に於いて(1)は人間の声帯振動の周期性を模倣して周期パルス音源信号を発生する有声音源発生回路、(2)は人間の気管での乱流振動の非周期性を模倣して雑音音源信号を発生する無声音源発生回路である。(3)は上記有声音源発生回路(1)から得られる周期パルス音源信号又は無声音源発生回路(2)から得られる雑音音源信号を選択する選択スイッチである。(4)は該スイッチ(3)から得られる音源信号にその信号本来の振巾成分を付与する乗算器、(5)は人間の声道の音解特性を模倣したデジタルフィルタであり、上記乗算器(4)から得られる有音又は無声音源信号を濾波する事

に依つて、有声音又は無声音の音声波形信号が出力される。(6)は該デジタルフィルタ(5)から得られるデジタル値の音声波形信号をアナログ値に変換するD・A変換器、(7)は該D・A変換器(6)からの音声波形信号に基づいて、合成音声を発音するスピーカである。(8)はパラメータメモリであり、上記有声音源発生回路(1)での周期パルス音源信号の周期を設定すると共に上記選択スイッチ(3)での音源信号の選択を指示するピッチパラメータPと上記乗算器(4)での振巾成分を設定するアンブパラメータAと、上記デジタルフィルタ(5)のフィルタ特性を決定するLBP周波数パラメータ $f_1, f_2, \dots, f_8$ と、がフレーム周期毎に貯えられている。

従つて、無声音声を合成する時には、ピッチパラメータPの値は“0”となつており、選択スイッチ(3)はこの値“0”を検知して、無声音源発生回路(2)からの雑音音源信号を乗算器(4)に導入し、この音源信号にアンブパラメータAの値を乗算する事に依つて、無声音源信号を得る事になる。一

方、有声音声を合成する時にはピッチパラメータPの値は“0”でない数値に依つてピッチ周期を示しており、この値Pに依つて、有声音源発生回路(1)はその周期パルス音源信号の周期を設定すると共に、選択スイッチ(3)はピッチパラメータPの値が“0”でない事を検知して有声音源発生回路(1)からの周期パルス音源信号を乗算器(4)に導入し、この音源信号にアンブパラメータAの値を乗算する事に依り、有声音源信号を得る事になる。

斯して、得られた有音又は無声音源信号はLBPパラメータ $f_1, f_2, \dots, f_8$ にてフィルタ特性が制御されたデジタルフィルタ(5)にて濾波され、さらにD・A変換(6)されてスピーカ(7)にて有音又は無音の合成音声が発音される。

斯様な音声合成装置に於いては、そのパラメータメモリ(8)に貯えておくべきパラメータP、A、 $f_1, f_2, \dots, f_8$ を予め元音信号に基づいて、分析しておく必要がある。この内LBP周波数 $f_1, f_2, \dots, f_8$ の分析過程については1981年2月2日発行の「日経エレクトロニクス」の第128頁～第129頁に於いて、

158頁の記事「線スペクトル周波数をパラメータとした音声合成法とそのLBI化」に詳しい。

またピッチパラメータP、及びアンブパラメータAの分析については従来からのPAROOR方式にて用いられていた分析過程が採用されている。

特にピッチパラメータPの値を“0”とするか否かに依る有声音の判定条件については特公昭5

5-34953号に詳しく記載されている様に、元音信号から抽出した変形相関関数W(n)の最大値P<sub>0</sub>を検出し、これに1次のPAROOR係数K<sub>1</sub>に0.5を乗じた値0.5K<sub>1</sub>を加算した値、即ち、 $P_0 + 0.5K_1$ を求め、この値を特定の閾値と比較し、

$P_0 + 0.5K_1 \geq \text{閾値}$ の時、元音信号が有音

音であり、逆に

$P_0 + 0.5K_1 < \text{閾値}$ の時、元音信号が無音

であると判定していた。

しかしながら、LBP音声合成方式に於いて、その音声分析での有音無音の判定処理時に、PAROOR係数K<sub>1</sub>を用いなければならないばかり

か、従自己相関関数  $W(\tau)$  の最大値  $P_{00}$  を導出する  
 為に多大の計算処理を必要とする欠点があり、斯る  
 音声分析の実時間処理を阻害する不都合があつた。

#### Ⅱ 発明の目的

本発明は上述の点に鑑みて為され有聲無聲の判定  
 処理の簡略化を図った音声合成方式を提供する  
 ものである。

#### Ⅲ 発明の構成

本発明の音声合成方式は、LSPパラメータ  $f_1, f_2, \dots, f_N$  の各平均値  $\bar{f}_1, \bar{f}_2, \dots, \bar{f}_N$  及びその標準偏差  $\sigma_1, \sigma_2, \dots, \sigma_N$  を予  
 め求めておき、時刻  $T$  の LSPパラメータ  $f_1(T), f_2(T), \dots, f_N(T)$  を次々

$$P_1(T) = \frac{f_1(T) - \bar{f}_1}{\sigma_1}$$

に変換する事に依つて正規化した後、さらに重み  
 付け係数  $O_1, O_2, \dots, O_N$  を設定して

$$A(T) = \sum_{i=1}^N O_i P_i(T)$$

を算出し、

この値  $A(T)$  に基づいて有聲無聲の判定を行なう  
 ものである。

て格納している。0509…は第1乃至第8の正規化  
 回路であり、第1番目 ( $i=1, 2, \dots, 8$ ) の  
 正規化回路05では、上記LSP分析回路03からこ  
 の時に得られる第1番目のLSP周波数パラメ  
 タ  $f_1(T)$  を記憶部04の第1番目のLSP周波数パ  
 ラメータの平均値  $\bar{f}_1$  とその標準偏差  $\sigma_1$  とを用  
 いて、

$$P_1(T) = \frac{f_1(T) - \bar{f}_1}{\sigma_1}$$

で表わす値  $P_1(T)$  に変換する。即ち、この値  $P_1$   
 (T)はこのフレーム周期  $T$  での各LSP周波数パ  
 ラメータ  $f_1(T)$  がどの程度平均値からはずれている  
 かを正規化して表わしている事となる。0600…は  
 第1乃至第8の乗算器であり、第1番目の正規化  
 回路05からの値  $P_1(T)$  に第1番目の重み付け係数  
 $O_1$  を掛けて値  $O_1 P_1(T)$  を算出する。07は加算  
 器であり、上記第1乃至第8の乗算器0600…から  
 の値  $O_i P_i(T)$  の総和

$$A(T) = \sum_{i=1}^N O_i P_i(T)$$

を算出する。即ち、この値  $A(T)$  が大きい時には、  
 このフレーム周期  $T$  での各LSP周波数パラメ

#### Ⅳ 実施例

第2図に本発明の音声合成方式に於ける元音声  
 の分析装置の一実施例を示す。00は元音声を入力  
 するマイクロフォン、01は該マイクロフォン(1)か  
 らの音声信号をディジタル形式に変換するA・D  
 変換器である。02は該A・D変換器01から音声信  
 号をLSP分析するLSP分析回路であり、基本的  
 的には第1図の合成装置のディジタルフィルク(5)  
 とは逆のフィルタ特性をもち、フレーム周期  $T$  毎  
 にLSP周波数パラメータ  $f_1(T), f_2(T), \dots, f_8(T)$  を算出すると共に、音値信号に対応する残  
 差信号のエネルギー値としてのアンパパラメータ  
 $A(T)$  が出力される。03は上記A・D変換器01から  
 の音声信号のピーク値を保持するピークホールド  
 機能を備えたピッチ抽出回路であり、ピーク間の  
 時間をピッチ周期  $P$  として出力する。04は記憶部  
 であり、各LSPパラメータ  $f_1, f_2, \dots, f_8$   
 に関して、夫々の値の存在分布を予め求めて  
 おき、各値の平均値  $\bar{f}_1, \bar{f}_2, \dots, \bar{f}_8$  とその  
 標準偏差  $\sigma_1, \sigma_2, \dots, \sigma_8$  との各値を算出し

タ  $f_1(T)$  が標準的な値から大きく外れており、逆  
 にこの値  $A(T)$  が小さい時にはそれが標準的な値に  
 近いものとなつている事を示している。08は比較  
 回路であり、上記加算器07から得られる値  $A(T)$  を  
 特定の閾値  $V_{th}$  と比較する比較回路であり、  
 $A(T) > V_{th}$  の時、各LSP周波数パラメ  
 タ  $f_1(T)$  が標準的な値から大きく外れていると判  
 定され、即ち、無聲音声であると判定され、無音  
 音声検知信号が出力され、逆に

$A(T) \leq V_{th}$  の時、各LSP周波数パラメ  
 タ  $f_1(T)$  が標準的な値の範囲に収まると判定され  
 即ち有聲音声であると判定され有聲音声検知信号  
 が出力される。09は上記ピッチ抽出回路03からの  
 ピッチ周期  $P$  と等レシダ側からの値  $0$  とを  
 切換出力する切換スイッチであり、上記比較回路  
 08からの有聲音声検知信号を受けた時、ピッチ周  
 期  $P$  を出力し、逆に無聲音声検知信号を受けた時  
 値  $0$  を出力する構成となつている。即ち、こ  
 の切換スイッチ09からはピッチパラメータが出力  
 される事となる。

斯して得られた各パラメータ $f_1, f_2, \dots, f_8, A, P$ を第1図のパラメータメモリ(8)に格納すれば、ピッチパラメータ $P$ の値が「0」であるかピッチ周期 $P$ であるかに依つて選択スイッチ(3)は無声音源発生回路(2)からの雑音音源信号又は有声音源発生回路(1)の周期パルス音源信号を選択する事となる。

又、上述の如くして得られたパラメータ $f_1, f_2, \dots, f_8, A, P$ を第1図のパラメータメモリ(8)に格納する事なく、デジタルフィルタ(5)、乗算器(4)、並びに有声音源発生回路(1)及び選択スイッチ(3)に直接伝送すれば、実時間の分析合成が可能となる。

次に「お知らせします」なる音声の分析を例に挙げ、第3図を参照して有声音無声音の判定条件の具体例を示す。LBP周波数パラメータ $f_1, f_2, \dots, f_N$ の次数は $N=8$ であり、各重み付け係数は全て $G_i=1$  ( $i=1, 2, \dots, 8$ )とし、さらに閾値 $V_{th}=\frac{N}{2}=4$ とした場合、第3図に示す如く、 $A(t) > V_{th}$ となるのは「OBHIRA

OBHIRASU」の音声の内、無声音である「BH」に対応する両時間領域 $T_a, T_c$ と、同じく無声音である「B」に対応する両時間領域 $T_b, T_d$ となつている。逆にこれ等の時間領域以外は無声音の時間領域となつている。又、これ等の時間領域 $T_a, T_b, T_c, T_d$ は従来方式の変形相関関数のピーク値 $P_m$ と1次のPARCOR係数 $K_1$ とを用いた無声音の判定領域 $T_a, T_b, T_c, T_d$ とほぼ一致している事が分かる。

尚、上述の実施例に於いて、音声分析装置の記憶部54に記憶しておくLBP周波数パラメータの平均値 $\bar{f}_1, \bar{f}_2, \dots, \bar{f}_8$ 及びその標準偏差 $\sigma_1, \sigma_2, \dots, \sigma_8$ としては有声音及び無声音のあらゆる音声について抽出したLBP周波数パラメータ $f_1, f_2, \dots, f_8$ に基づいて、導出したものを用いても良いが、これを有声音のみのあらゆる音声について抽出したパラメータ $f_1, f_2, \dots, f_8$ に基づいて導出すれば、この為の演算処理が軽減できる上に、より正確な有声音無声音の判定処理が望める。

#### ㌦ 発明の効果

本発明の音声合成方式は以上の説明から明らか如く、LBP方式を採用して、そのLBP周波数パラメータからなる値A(t)を用いて有声音無声音の判定を行なう事ができるので、多大な計算処理を必要とする自己相関関数の最大値を用いた従来方式に比べて有声音無声音の判定処理の為の計算費の大巾な低減が図れる。従つて有声音無声音の判定処理を簡略化でき、実時間の音声分析処理が望める。

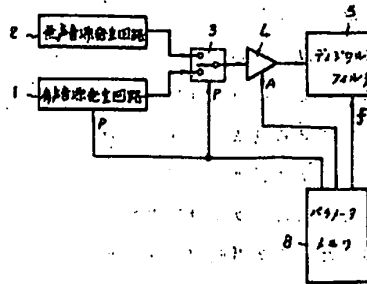
#### 4. 図面の簡単な説明

第1図はLBP方式を採用した音声合成装置の基本構成図、第2図は本発明の音声合成方式に用いられる音声分析装置の構成図、第3図は本発明方式に用いられる値A(t)の時間曲線図である。

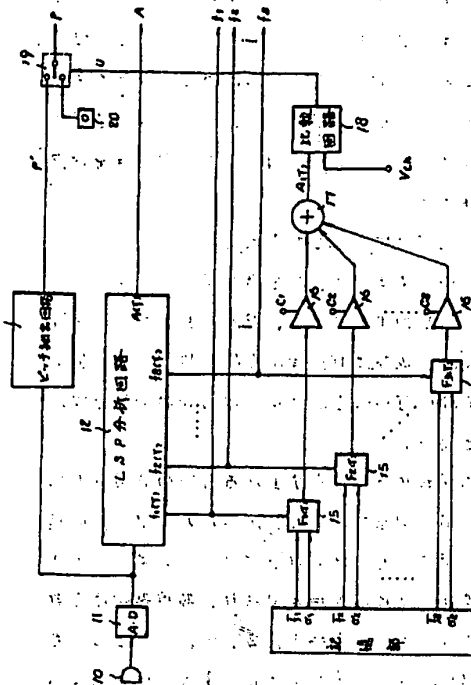
(1)…有声音源発生回路、(2)…無声音源発生回路、(3)…選択スイッチ、(4)…デジタルフィルタ、(5)…乗算器、(6)…パラメータメモリ、(7)…LBP分析回路、(8)…記憶部、(9)…正規化回路、(10)…乗算器、(11)…加算器、(12)…比較回路、(13)…切換スイッチ。

BEST AVAILABLE COPY

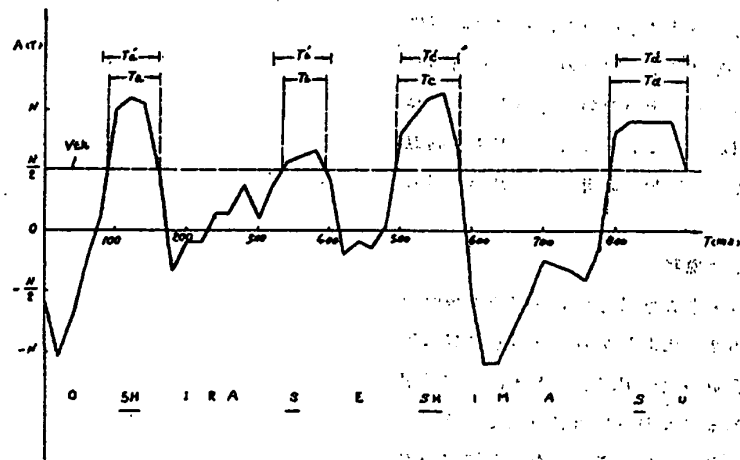
第1圖



第2圖



第3圖



BLANK PAGE